COMBINATORIAL PROBLEMS IN FINITE FIELDS AND SIDON SETS

JAVIER CILLERUELO

Abstract. We use Sidon sets to present an elementary method to study some combinatorial problems in finite fields. We obtain classic and more recent results avoiding the use of exponential sums, the usual tool to deal with these problems.

1. Introduction

Sárközy ([12], [13]) proved the solubility of the equations $x_1x_2 + x_3x_4 = 1$ and $x_1x_2 = x_3 + x_4,\ x_i \in A_i$ for arbitrary sets $A_i \subset \mathbb{F}_p$ when $|A_1||A_2||A_3||A_4| \gg p^3$. It was extended to any field $\mathbb{F}_q$ in [8]. The proof is based in estimates of exponential sums and they ask in problem 3 of [2] for an elementary algebraic proof of the solubility of these equations.

Many others combinatorial problems in $\mathbb{F}_q$ have been studied in recent years: incidence problems, sum-product estimates, distribution of small powers of a generator.

We present an elementary method to study these kind of problems. Our method is simple, combinatorial and avoids the use of exponential sums, the usual tool to deal with these problems. Beside quick and elementary proofs of known results we also provide some new results.

The main tool in our approach are Sidon sets, which are important objects in combinatorial number theory.

2. Sidon sets

Let G be a finite abelian group. For any sets $A, B \subset G$ and $x \in G$, we write $r_{A-B}(x)$ for the number of representations of $x = a - b,\ a \in A,\ b \in B$ and we have the familiar identities

$$\sum_{x\in G} r_{A-B}(x) = |A||B| \tag{2.1}$$

$$\sum_{x\in G} r^2_{A-B}(x) = \sum_{x\in G} r_{A-A}(x) r_{B-B}(x). \tag{2.2}$$

Definition 1. *We say that a set $\mathcal{A} \subset G$ is a Sidon set if $r_{\mathcal{A}-\mathcal{A}}(x) \le 1$ whenever $x \neq 0$.*

By counting the number of differences $a - a'$, we can see that $|\mathcal{A}| < |G|^{1/2} + 1/2$ when $\mathcal{A}$ is a Sidon set. The most interesting Sidon sets are those with large cardinality. In other words, those with $|\mathcal{A}| = \sqrt{|G|} - \delta$ where δ is a small number. As usual we write $\delta_+ = \max(0, \delta)$.

We state our main theorem.

Theorem 2.1. *Let $\mathcal{A}$ be a Sidon set in a finite abelian group G with $|\mathcal{A}| = \sqrt{|G|} - \delta$. Then, for all $B, B' \subset G$ we have*

$$\#\{(b, b') \in B \times B',\ b + b' \in \mathcal{A}\} = \frac{|\mathcal{A}|}{|G|}|B||B'| + \theta(|B||B'|)^{1/2}|G|^{1/4}, \tag{2.3}$$

with $|\theta| < 1 + \frac{|B|}{|G|}\delta_+$.

Proof. Since $\mathcal{A}$ is a Sidon set then

$$\begin{aligned}\sum_{x\in G} r_{B-B}(x)r_{\mathcal{A}-\mathcal{A}}(x) &= |\mathcal{A}||B|+\sum_{x\neq 0} r_{B-B}(x)r_{\mathcal{A}-\mathcal{A}}(x)\\ &\le |\mathcal{A}||B|+\sum_{x\neq 0} r_{B-B}(x)=|\mathcal{A}||B|+|B|^2-|B|.\end{aligned}\tag{2.4}$$

Using this inequality and identities (2.1) and (2.2) we have

$$\begin{aligned}\sum_{x\in G}\left(r_{\mathcal{A}-B}(x)-\frac{|\mathcal{A}||B|}{|G|}\right)^2 &\le \sum_{x\in G} r_{B-B}(x)r_{\mathcal{A}-\mathcal{A}}(x)-\frac{|\mathcal{A}|^2|B|^2}{|G|}\\ &\le |B|(|\mathcal{A}|-1)+|B|^2\frac{|G|-|\mathcal{A}|^2}{|G|}.\end{aligned}\tag{2.5}$$

We observe that

$$\#\{(b,b')\in B\times B',\ b+b'\in\mathcal{A}\}-\frac{|B||B'||\mathcal{A}|}{|G|}=\sum_{b'\in B'}\left(r_{\mathcal{A}-B}(b')-\frac{|\mathcal{A}||B|}{|G|}\right).$$

To finish the proof we first apply Cauchy's inequality, then inequality (2.5) and finally substitute $|\mathcal{A}|=|G|^{1/2}-\delta$.

$$\begin{aligned}\left|\sum_{b'\in B'}\left(r_{\mathcal{A}-B}(b')-\frac{|\mathcal{A}||B|}{|G|}\right)\right|^2 &\le |B'|\left(|B|(|\mathcal{A}|-1)+|B|^2\frac{|G|-|\mathcal{A}|^2}{|G|}\right)\\ &= |B'||B|\left(|G|^{1/2}-\delta-1+|B|\frac{\delta(2|G|^{1/2}-\delta)}{|G|}\right)\\ &< |B||B'||G|^{1/2}\left(1+2\max(0,\delta)\frac{|B|}{|G|}\right).\end{aligned}$$

□

The Sidon sets we will consider satisfy $\delta\le 1$ and $|B|=o(|G|)$, so $|\theta|\le 1+o(1)$.

2.1. **Examples of dense Sidon sets.**

The three families of Sidon sets we will describe next, have maximal cardinality in their ambient group G. Let g be a generator of $\mathbb{F}_q$.

Example 1. *Let $p(x),r(x)$ be polynomials of degree ≤ 2 in $\mathbb{F}_q[X]$ such that $p(x)-\mu r(x)$ is not a constant for any $\mu\in\mathbb{F}_q$. The set*

$$\mathcal{A}=\{(p(x),r(x)):\ x\in\mathbb{F}_q\}$$

is a Sidon set in $\mathbb{F}_q\times\mathbb{F}_q$. A special case is the set $\mathcal{A}=\{(x,x^2):\ x\in\mathbb{F}_q\}$.

We have to check that the relation $(p(x_1),r(x_1))-(p(x_2),r(x_2))=(e_1,e_2)$ determines x_1 and x_2 when $(e_1,e_2)\neq(0,0)$. If $p(x)$ is linear then from $p(x_1)-p(x_2)=e_1$ we obtain $x_1=x_2+\lambda$ for some λ. Thus, $r(x_2+\lambda)-r(x_2)=e_2$ is a linear equation and we obtain x_2 and then x_1. If $p(x)$ is quadratic we consider μ such that $p(x)-\mu r(x)$ is a linear polynomial and we proceed as above.

Example 2. *For any generator g of $\mathbb{F}_q^*$, the set*

$$\mathcal{A}=\{(x,g^x):\ x\in\mathbb{Z}_{q-1}\}\tag{2.6}$$

is a Sidon set in $\mathbb{Z}_{q-1}\times\mathbb{F}_q$.

From $(x_1,g^{x_1})-(x_2,g^{x_2})=(e_1,e_2)\neq(0,0)$ we have $x_1-x_2\equiv e_1 \pmod{q-1}$ and hence $g^{x_1}=g^{e_1+x_2}$. Putting this in $g^{x_1}-g^{x_2}=e_2$ we get $g^{x_2}(g^{e_1}-1)=g^{e_2}$.

If $e_1=0$ then $e_2=0$ but we have assumed that $(e_1,e_2)\neq(0,0)$. If $e_1\neq 0$ the last equality determines x_2, and then x_1.

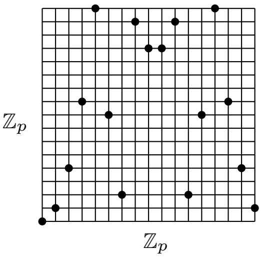
Example 3. *For any pair of generators g_1, g_2 of $\mathbb{F}_q^*$, the set*

$$\mathcal{A} = \{(x,y) \in \mathbb{Z}_{q-1} \times \mathbb{Z}_{q-1} :\ g_1^x + g_2^y = 1\} \tag{2.7}$$

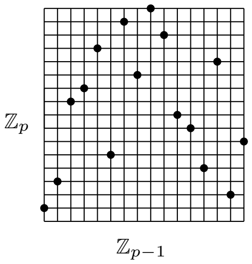
is a Sidon set in $\mathbb{Z}_{q-1} \times \mathbb{Z}_{q-1}$. Since translations preserve Sidoness property, for any $\lambda \neq 0$, the sets $\mathcal{A} = \{(x,y) :\ g_1^x + g_2^y = \lambda\}$ and $\mathcal{A} = \{(x,y) :\ g_1^x - g_2^y = \lambda\}$ are also Sidon sets.

To see that $\mathcal{A}$ is a Sidon set we have to prove that $(x_1,y_1)-(x_2,y_2) = (e_1,e_2),\ (e_1,e_2) \neq (0,0)$ determines x_1, x_2 under the conditions $g_1^{x_1} + g_2^{y_1} = g_1^{x_2} + g_2^{y_2} = 1$ in $\mathbb{F}_q$. We observe that $x_1 - x_2 \equiv e_1 \pmod{q-1}$ and $y_1 - y_2 \equiv e_2 \pmod{q-1}$ imply that $g_1^{x_1} = g_1^{x_2+e_1}$ and $g_2^{y_1} = g_2^{y_2+e_2}$ in $\mathbb{F}_q$ and we obtain $g_1^{x_2+e_1} + g_2^{y_2+e_2} = g_1^{x_2} + g_2^{y_2} = 1$ in $\mathbb{F}_q$. Thus $(1-g_2^{y_2})g_1^{e_1} + g_2^{y_2+e_2} = 1$ and we obtain y_2 and then x_2, x_1 and y_1.

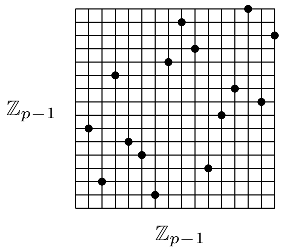
When q is a prime p we can identify $\mathbb{F}_p$ with $\mathbb{Z}_p$. We ilustrate in the pictures below the three examples of Sidon sets described above.



Example 1



Example 2



Example 3

The Sidon sets above, with $q, q-1$ and $q-2$ elements respectively, have maximal cardinality in their ambient groups. The values of $\delta = |G|^{1/2} - |\mathcal{A}|$ are $\delta = 0, 1/2 - o(1)$ and 1 respectively.

3. Equations in $\mathbb{F}_q$

The strategy is to choose the Sidon set $\mathcal{A}$ and the sets B and B' according with the equation we want to study. We start with the easiest example which, however, we have not seen in the literature.

Theorem 3.1. *For any $x \in \mathbb{F}_q$ let $X(x), Y(x)$ be any pair of subsets of $\mathbb{F}_q$ and put $T = \Big(\sum_x |X(x)|\Big)\Big(\sum_x |Y(x)|\Big)$. Then, the number of solutions S of*

$$x' + y' = (x+y)^2, \quad x' \in X(x),\ y' \in Y(y)$$

is

$$S = \frac{T}{q} + \theta\sqrt{qT}, \qquad |\theta| \leq 1. \tag{3.1}$$

Proof. We consider the Sidon set $\mathcal{A} = \{(x,x^2) :\ x \in \mathbb{F}_q\}$ and the sets $B = \{(x,x') : x' \in X(x)\}$ and $B' = \{(y,y') :\ y' \in Y(y)\}$. We observe that $(x,x') + (y,y') \in \mathcal{A} \iff x' + y' = (x+y)^2$. Thus $S = |\{(b,b') \in B \times B' :\ b + b' \in \mathcal{A}\}|$ and we apply Theorem 2.1. □

Corollary 3.1. *Let $A_1, A_2, A_3, A_4 \subset \mathbb{F}_q$ and put $T = |A_1||A_2||A_3||A_4|$. Then, the number of solutions of the equation*

$$x_1 + x_2 = (x_3 + x_4)^2,\ x_i \in A_i \tag{3.2}$$

is

$$S=\frac{T}{q}+\theta\sqrt{qT},\qquad |\theta|\le 1. \tag{3.3}$$

In particular, the number of solutions of

$$x_1+x_2=z^2,\quad x_1\in A_1,\ x_2\in A_2,\ z\in\mathbb{F}_q \tag{3.4}$$

is

$$|A_1||A_2|+\theta\sqrt{|A_1||A_2|q},\qquad |\theta|\le 1. \tag{3.5}$$

Proof. Take $X(x)=\begin{cases}A_1, & x\in A_3\\ \emptyset \text{ otherwise}\end{cases}$ and $Y(x)=\begin{cases}A_2, & x\in A_4\\ \emptyset \text{ otherwise}\end{cases}$ in Theorem 3.1. For the second part of the corollary, we observe that each solution of (3.4) corresponds to exactly q solutions of (3.2) when we take $X_3=X_4=\mathbb{F}_q$. □

Shkredov [14] used Weil's bound for exponential sums with multiplicative characters to prove the following theorem for $q=p$ and the condition $|X_1||X_2|>20p$.

Theorem 3.2. *Let $X_1,X_2\subset\mathbb{F}_q$ with $|X_1||X_2|>2q$. Then there exist $x,y\in\mathbb{F}_q$ such that $x+y\in X_1$ and $xy\in X_2$.*

Proof. Actually, we will estimate the number of such pairs (x,y), which is the number of solutions of the equation

$$(x_1/2-z)(x_1/2+z)=x_2,\qquad x_1\in X_1,\ x_2\in X_2,\ z\in\mathbb{F}_q.$$

We observe that this equation is equivalent to the equation $(x_1/2)^2-x_2=z^2$. In order to apply (3.4) we split $X_1=X_{11}\cup X_{12}$ in such a way that the squares in each set are all distinct and we apply (3.5) separately to $A_1=\{x_1^2/2:\ x_i\in X_{11}\},\ A_2=-X_2$ and to $A_1=\{x_1^2/2:\ x_i\in X_{12}\},\ A_2=-X_2$. Then we obtain that the number of solutions of the equation $(x_1/2)^2-x_2=z^2,\quad x_1\in X_1,\ x_2\in X_2,\ z\in\mathbb{F}_q$ is $|X_{11}||X_2|+\theta_1\sqrt{|X_{11}||X_2|q}+|X_{12}||X_2|+\theta_2\sqrt{|X_{12}||X_2|q}=|X_1||X_2|+\theta\sqrt{|X_1||X_2|q},\quad |\theta|\le\sqrt{2}$. Finally we observe that this number is positive if $|X_1||X_2|>2q$. □

In [12] and [13] Sárközy studied the number of solutions of the congruences $x_1x_2-x_3x_4\equiv\lambda \pmod p$ and $x_1x_2-x_3-x_4\equiv\lambda \pmod p$, $x_i\in X_i$, using exponential sums. In [8] these congruences were generalized to equations in finite fields. We show how they can be deduced quickly as a consequence of Theorem 2.1.

Theorem 3.3. *For any $x\in\mathbb{F}_q^*$ let $X(x),Y(x)$ be any pair of subsets of $\mathbb{F}_q$ and put $T=\left(\sum_x|X(x)|\right)\left(\sum_x|Y(x)|\right)$. Then, the number of solutions S of the equation*

$$x'+y'=xy,\quad x'\in X(x),\ y'\in Y(y)$$

is

$$S=\frac{T}{q}+\theta\sqrt{qT},\qquad |\theta|\le 1+o(1). \tag{3.6}$$

Proof. We consider the Sidon set $\mathcal{A}=\{(x,g^x):\ x\in\mathbb{Z}_{q-1}\}$ and the sets $B=\{(\log x,x'):\ x'\in X(x)\}$ and $B'=\{(\log y,y'):\ y'\in Y(y)\}$. We observe that $(\log x,x')+(\log y,y')\in\mathcal{A}\iff x'+y'=g^{\log x+\log y}=xy$. Thus $S=|\{(b,b')\in B\times B':\ b+b'\in\mathcal{A}\}|$ and then we apply Theorem 2.1. □

Corollary 3.2. *Let $X_1,X_2\subset\mathbb{F}_q^*$ and $X_3,X_4\subset\mathbb{F}_q$ and put $T=|X_1||X_2||X_3||X_4|$. The number of solutions of $x_1x_2=x_3+x_4,\ x_i\in X_i$ is*

$$S=\frac{T}{q}+\theta\sqrt{Tq},\qquad |\theta|\le 1+o(1).$$

Proof. We take $X(x)$ and $Y(y)$ as in Corollary 3.1 and use them in Theorem 3.3. □

Corollary 3.3. *Let $X_1, X_2 \subset \mathbb{F}_q^*$ and $X_3, X_4 \subset \mathbb{F}_q$ and put $T = |X_1||X_2||X_3||X_4|$. The number of solutions S of $x_2x_3 - x_1x_4 = 1,\ x_i \in X_i$ is*

$$S = \frac{T}{q} + \theta\sqrt{Tq}, \qquad |\theta| \le 1 + o(1).$$

Proof. We take $X(x) = \begin{cases} xX_3, \ x \in X_1^{-1} \\ \emptyset \text{ otherwise} \end{cases}$ and $Y(y) = \begin{cases} -yX_4, \ y \in X_2^{-1} \\ \emptyset \text{ otherewise} \end{cases}$ in Theorem 3.3. In this way we obtain the number of solutions of the equation $x_1^{-1}x_2^{-1} = x_1^{-1}x_3 - x_2^{-1}x_4$, which is equivalent to the equation of the corollary. □

Corollary 3.4. *Let $A_1, A_2 \in F_q$. Then the number of solutions of*

$$1 + x_1x_2 = z^2, \qquad x_i \in A_i,\ z \in \mathbb{F}_q^* \tag{3.7}$$

is $|A_1||A_2| + \theta\sqrt{|A_1||A_2|q}$ for some $|\theta| \le 4$.

Proof. We observe that there are two solutions of (3.7) for each solution of $1 + x_1x_2 = r,\ r \in R = \{x^2 :\ x \ne 0\}$. Also we observe that each $z \in R$ can be written in exactly $|R|$ ways as $z = z_1z_2,\ z_1, z_2 \in R$. We will estimate the number of solutions of $1 + x_1x_2 = z_1z_2,\ x_1 \in A_1,\ x_2 \in A_2,\ z_1, z_2 \in R$. Corollary 3.3 gives that this number is $\frac{|A_1||A_2||R|^2}{q} + \theta|R|\sqrt{|A_1||A_2|q}$. Thus, the number of the solutions of (3.7) is

$$\frac{2}{|R|}\left(\frac{|A_1||A_2||R|^2}{q} + \theta|R|\sqrt{|A_1||A_2|q}\right).$$

□

Theorem 3.4. *For any $x \in \mathbb{F}_q^*$, let $X(x), Y(x)$ be any pair of subsets of $\mathbb{F}_q^*$ and put $T = \left(\sum_x |X(x)|\right)\left(\sum_x |Y(x)|\right)$. The number of solutions of $xy - x'y' = 1,\ x' \in X(x),\ y' \in Y(y)$ is*

$$S = \frac{T}{q} + \theta\sqrt{Tq}, \qquad |\theta| \le 1 + o(1).$$

Proof. We consider the Sidon set $\mathcal{A} = \{(x,y) : g^x - g^y = 1\} \subset \mathbb{Z}_{q-1} \times \mathbb{Z}_{q-1}$ and the sets $B = \{(\log x, \log x') :\ x' \in X(x)\}$ and $B' = \{(\log y, \log y') :\ y' \in Y(y)\}$. It is clear that $S = |\{(b, b') \in B \times B' :\ b + b' \in \mathcal{A}\}|$. Now we apply Theorem 2.1. □

We observe that this theorem also gives an alternative proof of Corollary 3.3 by taking $X(x)$ and $Y(y)$ as in Corollary 3.1.

4. Sum-product estimates

Sometimes we are interested in estimating the number of elements of a Sidon set in a set B. Of course, for an arbitrary set B we cannot say anything, since we can find large sets B for which $|\mathcal{A} \cap B| = 0$ and other sets for which $|\mathcal{A} \cap B| = |\mathcal{A}|$. However we get the following lemma which is useful in some situations.

Lemma 4.1. *Let $\mathcal{A}$ be a Sidon set in G with $|\mathcal{A}| = |G|^{1/2} - \delta$. For any $B, B' \subset G$ we have*

$$|\mathcal{A} \cap B| \le \frac{|B + B'||\mathcal{A}|}{|G|} + \theta\left(\frac{|B + B'|}{|B'|}\right)^{1/2}|G|^{1/4}, \tag{4.1}$$

with $|\theta| \le 1 + \max(0, \delta)\frac{|B'|}{|G|}$.

Proof. As a consequence of Theorem 2.1 we have

$$\begin{aligned}|B'||\mathcal{A}\cap B| &= |\{(-b',b+b'):\ b\in B,\quad b'\in B',\ -b'+(b+b')\in\mathcal{A}\}|\\ &\le \{(b',b''):\ b'\in(-B')\times(B+B'),\ b'+b''\in\mathcal{A}\}|\\ &\le \frac{|\mathcal{A}||B'||B+B'|}{|G|}+\theta\sqrt{|B'||B+B'|}|G|^{1/4}.\end{aligned}$$

□

The lemma above gives quick proofs of some sum-product estimates obtained in recent years.

Theorem 4.1 (Garaev [3]). *Let $A_1, A_2\subset\mathbb{F}_q^*$ and $A_3\subset\mathbb{F}_q$. We have*

$$\max(|A_1A_2|,|A_1+A_3|)\ge C\min\left(\sqrt{|A_1|q},\sqrt{|A_1|^2|A_2||A_3|/q}\right)$$

with $C=\frac{\sqrt{5}-1}{2}+O(\frac{|A_2||A_3|}{q^2})$.

Proof. Take $\mathcal{A}=\{(x,g^x):\ x\in\mathbb{Z}_{q-1}\}$, $B=(\log A_1)\times A_1$ and $B'=(\log A_2)\times A_3$. Since all the elements $(\log a_1,a_1)$ are in $\mathcal{A}$ we have that $|\mathcal{A}\cap B|=|A_1|$. On the other hand we observe that $|B+B'|=|A_1A_2||A_1+A_3|$. Lemma 4.1 implies the inequality

$$|A_1|\le\frac{|A_1A_2||A_1+A_3|}{q}+\theta\sqrt{q\frac{|A_1A_1||A_1+A_3|}{|A_2||A_3|}}\tag{4.2}$$

with $|\theta|\le 1+\frac{|A_2||A_3|}{q(q-1)}$.

Let $\alpha=\frac{2+\theta^2-\sqrt{4\theta^2+\theta^4}}{2}=(\frac{\sqrt{5}-1}{2})^2+O(\frac{|A_2||A_3|}{q^2})$ be the smallest solution of $\alpha=\frac{(1-\alpha)^2}{\theta^2}$. If $\frac{|A_1A_1||A_1+A_3|}{q}<\alpha|A_1|$, inequality (4.2) implies that $|A_1A_2||A_1+A_3|>\frac{(1-\alpha)^2}{\theta^2}\frac{|A_1|^2|A_2||A_3|}{q}=\alpha\frac{|A_1|^2|A_2||A_3|}{q}$. Thus, $|A_1A_2||A_1+A_3|\ge\alpha\min(|A_1|q,\frac{|A_1|^2|A_2|A_3|}{q})$. □

We can mimic the proof above to get the following sum-product estimates.

Theorem 4.2 (Garaev-Shen [6]). *Let $A_1,A_2,A_3\subset\mathbb{F}_q^*$. We have*

$$\max(|(A_1+1)A_2|,|A_1A_3|)\ge C\min\left(\sqrt{|A_1|q},\sqrt{|A_1|^2|A_2||A_3|/q}\right)$$

with $C=\frac{\sqrt{5}-1}{2}+O(\frac{|A_2||A_3|}{q^2})$.

Proof. Consider the Sidon set $\mathcal{A}=\{(x,y):\ g^x-g^y=1\}$, the sets $B=\log(A_1+1)\times\log A_1$ and $B'=\log A_2\times\log A_3$ and proceed as in the former proof. □

Theorem 4.3 (Solymosi [15], Hart-Li-Shen [9]). *Let $p(x),q(x)$ be polynomials in $F_q[X]$ of degree ≤ 2 such that $p(x)-\mu q(x)$ is not a constant for any $\mu\in\mathbb{F}_q$. For any $A_1,A_2,A_3\subset\mathbb{F}_q$ we have*

$$\max(|p(A_1)+A_2|,|q(A_1)+A_3|)\ge\frac{\sqrt{5}-1}{2}\min\left(\sqrt{|A_1|q},\sqrt{|A_1|^2|A_2||A_3|/q}\right).$$

Proof. Consider the Sidon set $\mathcal{A}=\{(p(x),q(x)):\ x\in\mathbb{F}_q\}$, the sets $B=p(A_1)\times q(A_1)$ and $B'=A_2\times A_3$ and proceed as above. We observe that in this case $|\theta|\le 1$ and we can remove the error term in the constant. □

Indeed Solymosi proved that if $\{(x,f(x)):\ x\in\mathbb{F}_q\}\subset\mathbb{F}_q\times\mathbb{F}_q$ is a Sidon set then $\max(|A+A|,|f(A)+f(A)|)\gg\min(\sqrt{|A|q},|A|^2/\sqrt{q})$, but it can be proved that this set is a Sidon set if and only if $f(x)$ is a quadratic polynomial.

It should be noticed that our method only works for sum-product estimates for large sets. The study of sum-product estimates in finite fields for small sets was initiated in the seminal paper [1] and we don't see how to apply our method in these cases.

5. Distribution of Sidon sets in regular sets and applications

Many problems can be described by giving an asymptotic estimate of $|\mathcal{A}\cap B|$ where $\mathcal{A}$ is a Sidon set and B is a suitable set. For example, if $\mathcal{A}=\{(x,x^2):\ x\in\mathbb{Z}_p\}$ and $B=\mathbb{Z}_p\times I$, the quantity $|\mathcal{A}\cap B|$ counts two times the number of quadratic residues modulo p lying in I. We will see other examples in this section.

The expected number for $|\mathcal{A}\cap B|$ when B is a regular set is $|B||\mathcal{A}|/|G|$. Thus we write $E_{\mathcal{A}}(B)=|\mathcal{A}\cap B|-\frac{|B||\mathcal{A}|}{|G|}$.

The next lemma and Lemma 4.1 will be the main tools to prove asymptotic estimates for $|\mathcal{A}\cap B|$ in some situations. For simplicity we consider only the three Sidon sets described in section §2.

Lemma 5.1. *Let $\mathcal{A}$ be one of the three Sidon sets described in section §2 and $B\subset G$. For any set $C\subset G$, there exists $c\in C$ such that*

$$|E_{\mathcal{A}}(B)|\le 2\Big(q\frac{|B|}{|C|}\Big)^{1/2}+|E_{\mathcal{A}}(B^c)|+|E_{\mathcal{A}}(B_c)|$$

where $B^c=B\setminus(B+c)$ and $B_c=(B+c)\setminus B$.

Proof.

$$E_{\mathcal{A}}(B)=|\mathcal{A}\cap B|-\frac{|\mathcal{A}||B|}{|G|}=\frac{1}{|C|}\sum_{c\in C}\Big(|\mathcal{A}\cap(B+c)|-\frac{|\mathcal{A}||B|}{|G|}\Big) \tag{5.1}$$

$$+\frac{1}{|C|}\sum_{c\in C}\Big(|\mathcal{A}\cap B|-|\mathcal{A}\cap(B+c)|\Big). \tag{5.2}$$

We observe that $\sum_{c\in C}\Big(|\mathcal{A}\cap(B+c)|-\frac{|\mathcal{A}||B|}{|G|}\Big)=|\{(b,c)\in B\times C:\ b+c\in\mathcal{A}\}|-\frac{|\mathcal{A}||B||C|}{|G|}$. Then we apply Theorem 2.1 to bound the first sum by $2\Big(q\frac{|B|}{|C|}\Big)^{1/2}$.

For the second sum we observe that $|\mathcal{A}\cap B|-|\mathcal{A}\cap(B+c)|=|\mathcal{A}\cap B_c|-|\mathcal{A}\cap B^c|$. Since $|B_c|=|B^c|$ we get

$$\begin{aligned}\left|\frac{1}{|C|}\sum_{c\in C}\Big(|\mathcal{A}\cap B|-|\mathcal{A}\cap(B+c)|\Big)\right| &= \left|\frac{1}{|C|}\sum_{c\in C}\Big(|\mathcal{A}\cap B_c|-\frac{|\mathcal{A}||B_c|}{|G|}+\frac{|\mathcal{A}||B^c|-\mathcal{A}\cap B^c|}{|G|}\right|\\ &\le \frac{1}{|C|}\sum_{c\in C}\Big(|E_{\mathcal{A}}(B_c)|+|E_{\mathcal{A}}(B^c)|\Big)\\ &\le \max_{c\in C}\Big(|E_{\mathcal{A}}(B_c)|+|E_{\mathcal{A}}(B^c)|\Big).\end{aligned}$$

□

In the special case when B is a subgroup we can take $C=B$ and then $B^c=B_c=\emptyset$ for any $c\in C$. Thus, in this case we have

$$|E_{\mathcal{A}}(B)|\ll q^{1/2}.$$

We obtain as a corollary a well known result about the Fermat equation in finite fields.

Corollary 5.1. *Let Q,Q' be subgroups of $\mathbb{F}_q^*$. We have*

$$|\{(x,y)\in Q\times Q':\ x+y=1\}|=\frac{|Q||Q'|}{q}+O(\sqrt{q}).$$

In particular, if $p\gg(rs)^2$ the Fermat congruence $x^r+y^s\equiv 1 \pmod p$ has nontrivial solutions.

Proof. Consider the Sidon set $\mathcal{A} = \{(x, y) : \ g^x + g^y = 1\}$ and take $B = C = Q \times Q'$. □

In general, the strategy is to take a large set C such that $|B^c|$ and $|B_c|$ are small compared with $|B|$. This is possible when B has some kind of regularity (subgroups, cartesian product of arithmetic progressions, convex sets...). Then, we apply the lemma above again. We illustrate what we mean with an example.

Theorem 5.1. *Let $I, J \subset \mathbb{Z}_{p-1}$ be two intervals. For any positive integer r we have*

$$\{(x, y) \in I \times J : \ g^x - g^y \equiv \lambda \pmod p\} = \frac{|I||J|}{p} + \theta_r\Big(\Big(\frac{|I||J|}{p^{3/2}}\Big)^{1/r} + 1\Big)\sqrt{p},$$

with $|\theta_r| \le 4^r$.

Proof. We proceed by induction on r. We consider the Sidon set $\mathcal{A} = \{(x, y) : \ g^x - g^y = \lambda\} \subset \mathbb{Z}_{p-1} \times \mathbb{Z}_{p-1}$ and the set $B = I \times J$.

Lemma 4.1 applied to this case implies that

$$|\mathcal{A} \cap B| \le \frac{|B + B|}{p} + 2\sqrt{p\frac{|B + B|}{|B|}} \le \frac{4|I||J|}{p} + 4\sqrt{p}.$$

Since $|E_{\mathcal{A}}(B)| \le \max(\frac{|B||\mathcal{A}|}{|G|}, |\mathcal{A} \cap B|)$, we have that $|E_{\mathcal{A}}(B)| \le \frac{4|I||J|}{p} + 4\sqrt{p}$, which proves Theorem 5.1 for $r = 1$.

We consider the auxiliar set $C = I' \times J'$ where $I' = \{0, \dots, \lfloor\alpha|I|\rfloor\}$ and $J' = \{0, \dots, \lfloor\alpha|J|\rfloor\}$ for a suitable α. We observe that $|C| \ge \alpha^2|I||J|$. Lemma 5.1 gives

$$|E_{\mathcal{A}}(B)| \le 2\frac{p^{1/2}}{\alpha} + |E_{\mathcal{A}}(B^c)| + |E_{\mathcal{A}}(B_c)|.$$

Now we observe that $B + c$ is a small translation of the rectangle $B = I \times J$. Thus we can write $B^c = B_1 \cup B_2$ and $B_c = B_3 \cup B_4$ where the sets B_i are rectangles with $|B_i| \le \alpha|I||J|$.

Thus,

$$|E_{\mathcal{A}}(B)| \le 2\frac{p^{1/2}}{\alpha} + |E_{\mathcal{A}}(B_1)| + |E_{\mathcal{A}}(B_2)| + |E_{\mathcal{A}}(B_3)| + |E_{\mathcal{A}}(B_4)|.$$

Assuming the statement of Theorem 5.1 for r and applying it for each B_i we get

$$|E_{\mathcal{A}}(B)| \le 2\frac{p^{1/2}}{\alpha} + 4\theta_r\Big(\Big(\frac{\alpha|I||J|}{p^{3/2}}\Big)^{1/r} + 1\Big)\sqrt{p}.$$

Taking $\alpha = \frac{1}{4^r}\left(\frac{p^{3/2}}{|I||J|}\right)^{1/(r+1)}$ we prove the statement of Theorem 5.1 for $r + 1$. □

It should be mentioned that, for the particular case $|I| = |J|$, Garaev [4] obtained the error term $O(|I|^{2/3}\log^{2/3}(|I|p^{-3/4} + 2) + p^{1/2})$. It can be checked that the error term in Theorem 5.1 is smaller than Garaev's error term. Actually, in the range $p^3 \ll |I||J| \ll p^3(\log p)^{\log\log p}$ it is smaller than the error term $O(p^{1/2}\log^2 p)$ established in [11].

For arbitrary intervals, Theorem 5.1 only gives $O_\epsilon(p^{1/2+\epsilon})$ for any $\epsilon > 0$, which is slightly weaker than $O(p^{1/2}\log^2 p)$, but our proof is elementary.

We can also get the same error term for similar problems as the problem of estimating $\{x \in I : \ x^2 \in J\}$ or $\{x \in I : \ g^x \in J\}$ by considering other Sidon sets.

5.1. **The difference set $\{g^x - g^y : \ 0, \le x, y \le L\}$.** Let g a primite root of $\mathbb{F}_p$. Many authors have studied the problem of determining the smallest number M such that $\{g^x - g^y : \ 0 \le x, y \le M\} = \mathbb{Z}_p$. Odlyzco has conjectured that it is posible to take $M \ll p^{1/2+\epsilon}$ but the exponent $3/4$ seems to be the natural barrier for this problem with the known methods.

From the result of Rudnick and Zaharescu [11] it follows that one can take any integer $M \ge c_0 p^{3/4}\log p$ where c_0 is a suitable constant. This result was improved to the range

$M > cp^{3/4}$ with $c = 10$ by M. Z. Garaev and K. L. Kueh [5], with $c = 4$ by S. V. Konyagin [10] and with $c = 2^{5/4}$ by V. García [7]. We improve these results.

Theorem 5.2. *Let g be a primitive root of $\mathbb{F}_p$. If p is large enough then*

$$\left\{g^x - g^y : \ 0 \le x, y < \sqrt{2}p^{3/4}\right\} = \mathbb{F}_p.$$

Proof. Suppose $\lambda \notin \{g^x - g^y : \ 0 \le x, y \le M\}$ and consider the Sidon set $\mathcal{A} = \{(x, y) : g^x - g^y = \lambda\}$. For each $t \ge 0$ we write

$$B_{\pm t} = [0, (M \pm t)/2]^2 \cup \left((\frac{p-1}{2}, \frac{p-1}{2}) + [0, (M \pm t)/2]^2 \right.$$

and we apply Theorem 2.1 with $B' = B = B_t$ (for a suitable t we will choose later) to obtain

$$S = |\{(b, b') \in B_t \times B_t, \ b + b' \in \mathcal{A}\}| \ge \frac{|\mathcal{A}||B_t|^2}{|G|} - \left(1 + \frac{|B_t|}{|G|}\right)|B_t||G|^{1/4}.$$

Thus

$$|B_t| \le \frac{|G|^{5/4}}{|\mathcal{A}| - |G|^{1/4}} \left(1 + \frac{|S|}{|G|^{1/4}|B_t|}\right).$$

Now we have to bound S. We also observe that $(x, y) \in \mathcal{A} \iff (y, x) + (\frac{p-1}{2}, \frac{p-1}{2}) \in \mathcal{A}$. Thus $\mathcal{A} \cap (B_0 + B_0) = \emptyset$.

Actually if we take $t = 0$ we have $S = 0, \ |B_0| \ge M^2/2$ and we can easily obtain that $M \le \sqrt{2}p^{3/4}(1 + o(1))$. To remove the $o(1)$ term we need to take a suitable $t > 0$.

We observe that $B_t + B_{-t} \subset B_0 + B_0$. Thus, if $b \in B_t, \ b' \in B_{-t}$ then $b + b' \notin \mathcal{A}$. After this observation we have

$$\begin{aligned} S &= |\{(b, b') : b, b' \in (B_t \setminus B_{-t}) : \ b + b' \in \mathcal{A}\}| \\ &= \sum_{b' \in (B_t \setminus B_{-t})} |(B_t \setminus B_t) \cap (\mathcal{A} - b')|. \end{aligned}$$

Notice that $\mathcal{A} - b'$ contains at most an element in each row and in each column. Since $B_t \setminus B_{-t}$ is the union of four rectangles of dimensions $t \times (M + t)/2$, we conclude that $|(B_t \setminus B_{-t}) \cap (\mathcal{A} - b')| \le 4t$ for any b'. Thus, $S \le 4t|B_t \setminus B_{-t}| \le 8t^2(M + t)$.

On the other hand we have $|B_t| \ge (M + t)^2/2$. Thus

$$(M + t)^2 \le \frac{2|G|^{5/4}}{|\mathcal{A}| - |G|^{1/4}} \left(1 + \frac{16t^2(M + t)}{|G|^{1/4}|B|}\right) \le \frac{2|G|^{5/4}}{|\mathcal{A}| - |G|^{1/4}} \left(1 + \frac{32t^2}{|G|^{1/4}(M + t)}\right)$$

and

$$M + t \le \left(\frac{2|G|^{5/4}}{|\mathcal{A}| - |G|^{1/4}}\right)^{1/2} \left(1 + \frac{16t^2}{|G|^{1/4}(M + t)}\right).$$

If $M < \sqrt{2}(p - 1)^{3/4}$ we are done. Otherwise $M \ge \sqrt{2}(p - 1)^{3/4} = \sqrt{2}|G|^{3/8}$. We write also $|\mathcal{A}| = |G|^{1/2} - 1$ and we choose $t = [|G|^{1/4}/32]$. Then, assuming that $|G|$ is large enough we have

$$\begin{aligned} M &\le \left(\frac{2|G|^{5/4}}{|G|^{1/2} - |G|^{1/4} - 1}\right)^{1/2} \left(1 + \frac{16t^2}{\sqrt{2}|G|^{5/8}}\right) - t \\ &\le \sqrt{2}|G|^{3/8} + \sqrt{2}|G|^{1/8} - \frac{|G|^{1/4}}{64} + \frac{65}{64} < \sqrt{2}|G|^{3/8} < \sqrt{2}p^{3/4}. \end{aligned}$$

□

6. A variant of the method

Suppose we want to study the distribution of the inverses of the elements of a small interval in $\mathbb{F}_p$. The set $\mathcal{A} = \{(x, x^{-1}) :\ x \in \mathbb{F}_p^*\} \subset \mathbb{F}_p \times \mathbb{F}_p$ is the natural set to deal with this problem. Unfortunately, this set is not a Sidon set and we cannot apply Theorem 2.1 directly.

The result we obtain in this case is weaker than what is obtained using Kloosterman sums. However our method is enough, for example, to prove that the set $\{x^{-1} :\ x \in I\}$ is well distributed in $\mathbb{F}_p$ when I is a not very small interval.

In the theorem below, using Kloosterman sums is possible to get the error term $O(\sqrt{|A_1||A_2||A_3||A_4|q})$ and the exponent $5/6$ can be substituted by $3/4$.

Theorem 6.1. *Let $A_1, A_2, A_3, A_4 \subset \mathbb{F}_q$. The number of solutions of*

$$(x_1 + x_2)(x_3 + x_4) = \lambda, \qquad x_i \in A_i$$

is

$$\frac{|A_1||A_2||A_3||A_4|}{q} + O\Big(\sqrt{|A_1||A_2||A_3||A_4|q}(1 + (|A_1||A_3|/q)^{1/4})\Big).$$

In particular, if I is an interval with $|I| \gg p^{5/6}$ we have $F_p^* \subset \{xy :\ x, y \in I\}$.

Proof. The proof is similar for any $\lambda \neq 0$. Thus we assume that $\lambda = 1$ and we consider the set $\mathcal{A} = \{(x, x^{-1}) :\ x \in \mathbb{F}_q^*\}$. We take $B = A_1 \times A_3,\ B' = A_2 \times A_4$ and we write $E = S - \frac{|B||B||\mathcal{A}|}{|G|}$. Following the first steps of the proof of Theorem 2.1 we get

$$E^2 \leq |B'| \left(|\mathcal{A}||B| + \sum_{x \neq 0} r_{\mathcal{A}-\mathcal{A}}(x) r_{B-B}(x) - \frac{|B|^2|\mathcal{A}|^2}{|G|} \right). \tag{6.1}$$

Now we observe that the sum is equal to $\sum_{b \neq b' \in B} r_{\mathcal{A}-\mathcal{A}}(b - b')$. Write $b = (a_1, a_3)$ and $b' = (a_1', a_3')$. Thus, the value of this sum is the number of solutions of

$$x - x' = a_1 - a_1', \quad x^{-1} - (x')^{-1} = a_3 - a_3',\ x \neq x',\ a_1, a_1' \in A_1,\ a_3, a_3' \in A_3$$

which is equivalent to the equation

$$1 - 4(a_1 - a_1')^{-1}(a_3 - a_3')^{-1} = (2x(a_1 - a_1')^{-1} - 1)^2.$$

Since x runs over all $x \neq 0$, the number of solutions is exactly the number of solutions of

$$1 - 4(a_1 - a_1')^{-1}(a_3 - a_3')^{-1} = z^2, \quad a_1, a_1' \in A_1,\ a_3, a_3' \in A_3\ z \in \mathbb{F}_q.$$

We have seen in Corollary 3.4 that the number of solutions of

$$1 - x_1 x_2 = z^2,\ x_1 \in X_1,\ x_2 \in X_2,\ z \in \mathbb{F}_q$$

is $|X_1||X_2| + \theta\sqrt{|X_1||X_2|q}$. For each a_1', a_3' we apply this to the sets $X_1 = \{4(a_1 - a_1')^{-1} :\ a_1 \in A_1 \setminus a_1'\}$ and $X_2 = \{(a_3 - a_3')^{-1} :\ a_3 \in A_3 \setminus a_3'\}$ and we obtain that the number of solutions of the equation is

$$\sum_{a_1' \in A_1,\ a_3 \in A_3'} \Big(|X_1||X_2| + \theta\sqrt{|X_1||X_2|q}\Big) \leq |A_1|^2|A_3|^2 + \theta(|A_1||A_3|)^{3/2}q^{1/2}.$$

Putting this in (6.1) we have

$$E^2 \leq |B'| \left((p-1)|B| + |A_1|^2|A_3|^2 + \theta(|A_1||A_3|)^{3/2}q^{1/2} - \frac{|\mathcal{A}|^2|B|^2}{|G|} \right).$$

Now we use that $|B'| = |A_2||A_4|,\ |B| = |A_1||A_3|,\ |\mathcal{A}| = p - 1,\ |G| = p^2$ to get

$$E^2 \ll |A_1||A_2||A_3||A_4|q\Big(1 + (|A_1||A_3|/q)^{1/2}\Big).$$

□

Acknowledgments: Parts of this work have been presented in several places: Conference on Arithmetic Combinatorics (IAS Princeton 2007), Primera Reunión Conjunta Sociedad Matemática Mexicana-Real Sociedad Matemática Española (Oaxaca, México, 2009) and Discrete Analysis Seminar (University of Cambridge, 2010). I would like to thank these institutions for his hospitality. Also I want to thank Igor Shparlinsky and M. Garaev for useful conversations about this work.

References

[1] Bourgain J., Katz N. and Tao T., A sum-product estimate in finite fields, and applications, *Geometric and Functional Analysis* 14 (2004) n1, 27–57.

[2] Csikvári P., Gyarmati K. and Sárközy A., Density and Ramsey type results on algebraic equations with restricted solution sets, *Journal of Combinatorial Theory, Series A*, to appear.

[3] Garaev M.Z., The sum-product estimate for large subsets of prime fields. *Proc. Amer. Math. Soc.* 136 (2008) n8, 2735–2739.

[4] Garaev M.Z., On the logarithmic factor in error term estimates in certain additive congruence problems, *Acta Arithmetica* 124 (2006) n1, 27–39.

[5] Garaev M.Z. and Kueh K., Distribution of special sequences modulo a large prime, *International Journal of Mathematics and Mathematical Sciences*, (2003) n 50, 3189–3194.

[6] Garaev M.Z. and Shen S., On the size of the set $A(A+1)$, *Mathematische Zeitschrift*, to appear

[7] García V.C. , A note on an additive problem with powers of a primitive root. *Bol. Soc. Mat. Mexicana* (3) 11 (2005) n1 1–4.

[8] Gyarmarti K. and Sárközy A., Equations in finite fields with restricted solution sets, II. (Algebraic equations.), *Acta Math. Hungar.* 119 (2008) n3, 259–280.

[9] Hart D., Li L., Shen C., Fourier analysis and expanding phenomena in finite fields. arXiv:0909.5471

[10] Konyagin S.V., Bounds of exponential sums over subgroups and Gauss sums, *Proc 4th Intern. Conf. Modern Problems of Number Theory and Its Applications*, Moscow Lomonosov State Univ., Moscow, (2002).

[11] Rudnik Z. and Zaharescu A., The distribution of spacing between small powers of a primitive root, *Israel Journal of Mathematics* 120 (2000)

[12] Sárközy A., On sums and products on residues modulo p. *Acta Arithmetica* 118 (2005) n4, 403–409.

[13] Sárközy A., On products and shifted products of residues modulo p. Proceedings of CANT 2005. *Integers 8* (2008) n2, 8pp.

[14] I. D. Shkredov, On monochromatic solutions of some nonlinear equations in $\mathbb{Z}/\mathbb{Z}_p$, arXiv:0909.3269

[15] J. Solymosi, Incidences and the spectra of graphs, *Combinatorial Number Theory and Additive Group Theory, Advanced Courses in Mathematics - CRM Barcelona* Birkhäuser Basel (2009)

[16] Vinh. Szemeredi-Trotter type theorem and sum-product estimate in finite fields, arXiv:0711.4427v1 [math.CO]